

Original document

DATA ABNORMALITY DETECTOR

Patent number: JP11338848

Publication date: 1999-12-10

Inventor: GOTO KENJI

Applicant: FFC KK; FUJI ELECTRIC CO LTD

Classification:

- international: G06F15/18; G06F15/18; G05B23/02; G08B25/00

- european:

Application number: JP19980144306 19980526

Priority number(s): JP19980144306 19980526

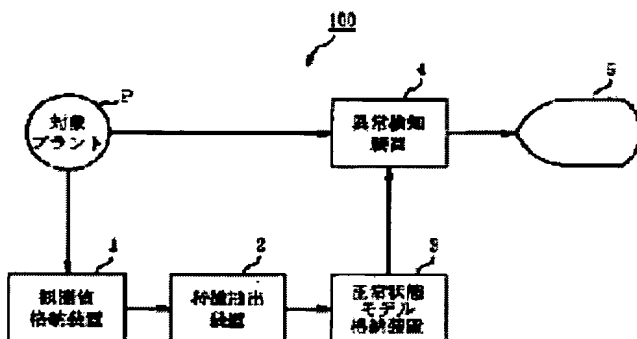
[View INPADOC patent family](#)

Report a data error here

Abstract of JP11338848

PROBLEM TO BE SOLVED: To easily and surely detect abnormality in unknown data.

SOLUTION: As data for monitor in a normal state of an object plant P, normal data are previously collected and based on these normal data, the feature of the normal data is extracted while using a self-organized map. Based on this feature, a feature map expressing the relation of a distance between respective output units is formed and stored as a normal state model. When unknown input data are inputted, based on these input data and the normal state model, an external distance dext expressing a distance between the input data and the weight vector of each output unit is found, and an internal distance dint is found as the minimum value of a distance between the output unit in this minimum external distance dext and the nearby output unit, and on the condition of $\text{internal distance dint} < \text{external distance dext}$, the possibility of abnormality in the input data is judged.



(19) 日本国特許庁 (J P)

(12) 公開特許公報 (A)

(11) 特許出願公開番号

特開平11-338848

(43) 公開日 平成11年(1999)12月10日

(51) Int.Cl. ⁶	識別記号	F I
G 0 6 F 15/18	5 5 0	G 0 6 F 15/18
	5 2 0	5 5 0 Z
G 0 5 B 23/02	3 0 2	5 2 0 E
G 0 8 B 25/00	5 1 0	G 0 5 B 23/02
		3 0 2 S
		G 0 8 B 25/00
		5 1 0 A

審査請求 未請求 請求項の数1 O L (全 6 頁)

(21) 出願番号 特願平10-144306

(22) 出願日 平成10年(1998) 5月26日

(71) 出願人 000237156

株式会社エフ・エフ・シー
東京都日野市富士町1番地

(71) 出願人 000005234

富士電機株式会社
神奈川県川崎市川崎区田辺新田1番1号

(72) 発明者 後藤 賢治

東京都日野市富士町1番地 株式会社エフ・エフ・シー内

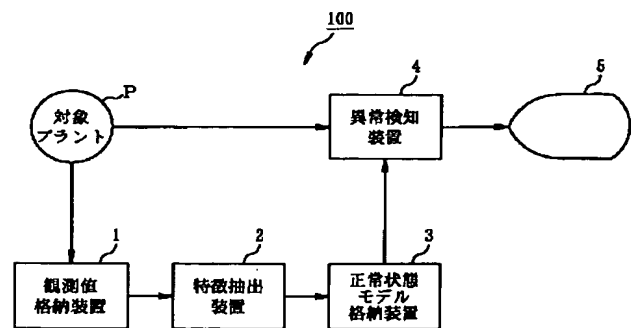
(74) 代理人 弁理士 森 哲也 (外3名)

(54) 【発明の名称】 データ異常検出装置

(57) 【要約】

【課題】 未知のデータの異常検出を容易確実に行う。

【解決手段】 対象プラントPが正常状態であるときの監視用データである正常データを予め収集しておき、この正常データをもとに自己組織化マップを用いて正常データの特徴を抽出し、これをもとに各出力ユニット間の距離関係を表す特徴マップを形成し正常状態モデルとして記憶しておく。未知の入力データを入力したときに、この入力データと正常状態モデルとをもとに、入力データと各出力ユニットの重みベクトルとの間の距離を表す外部距離 d_{ext} を求め、この外部距離 d_{ext} が最小となる出力ユニットとこの近傍の出力ユニットとの間の距離の最小値である内部距離 d_{int} とを求め、内部距離 $d_{int} < \text{外部距離 } d_{ext}$ となるときには、入力データは異常の可能性があると判断する。



【特許請求の範囲】

【請求項1】 予め収集した正常データをもとに、教師なし学習を行うニューラルネットワークを用いて正常データの特徴を抽出し、これを正常状態モデルとしてモデル記憶手段に格納する特徴抽出手段と、未知の入力データと前記正常状態モデルとに基づき、前記入力データを正常とみなすことの可能な正常判定基準値を動的に検出する正常判定基準値検出手段と、前記入力データと前記正常状態モデルとに基づき、前記入力データの異常判定のための判定値を演算する判定値演算手段と、前記正常判定基準値と前記判定値とに基づいて正常状態から乖離した異常状態を検出することにより前記入力データの異常検出を行う異常検出手段と、を備えることを特徴とするデータ異常検出装置。

【発明の詳細な説明】

【0001】

【発明の属する技術分野】この発明は、バッチプラント、連続系プラント等において、これらプラントにおける所定の監視用データをもとに対象プラントの異常検出を行うようにしたデータ異常検出装置に関する。

【0002】

【従来の技術】従来、バッチプラント、連続系プラント等において、異常予測、診断のような異常検出を行う場合等には、例えば監視対象プラントに設けたセンサ等から監視用データを収集してこれを分析し、例えば予め設定している異常判定用のしきい値と比較すること等によって、異常が発生しているかどうか、或いは異常が発生する兆候であるかどうか等を判断するようにしている。

【0003】

【発明が解決しようとする課題】しかしながら、上記従来の方法では、異常検出の対象となる監視用データを直接計測することは困難な場合が多い。そのため、通常の固定化したしきい値をもとに異常検出を行う方法では、確実に異常検出を行うために、各種センサ毎にしきい値を厳しく設定しているから、実際には異常でない場合でも異常として検出する場合があり、異常の検出回数が増加する。このために、オペレータは適切な判断を行うことが困難になり、異常の確認作業が増大し、誤確認によるミスが発生しやすくなるという問題がある。

【0004】これを回避するためには、互いに関連する監視用データを含めた総合的な判断方法の開発が必要となるが、関連する監視用データが把握されても、それぞれの値そのものの判断やこれら監視用データから対象プラントを総合的に判断する上において、明確な判断基準を定義することは難しい。

【0005】また、前述の異常検出用のしきい値を設定する際には、多種多様な異常状態に対する学習用のデータを予め収集して蓄積し、これを分析して作成したデータベースをもとにしているため、例えば事前に異常状態

のデータを用意することの困難な対象プラントに対しては、しきい値をもとに異常検出を行う方法を利用することができないという問題がある。

【0006】そこで、この発明は、上記従来の問題点に着目してなされたものであり、未知のデータの異常検出を容易確実に行うことの可能なデータ異常検出装置を提供することを目的としている。

【0007】

【課題を解決するための手段】上記目的を達成するために、本発明に係るデータ異常検出装置は、予め収集した正常データをもとに、教師なし学習を行うニューラルネットワークを用いて正常データの特徴を抽出し、これを正常状態モデルとしてモデル記憶手段に格納する特徴抽出手段と、未知の入力データと前記正常状態モデルとに基づき、前記入力データを正常とみなすことの可能な正常判定基準値を動的に検出する正常判定基準値検出手段と、前記入力データと前記正常状態モデルとに基づき、前記入力データの異常判定のための判定値を演算する判定値演算手段と、前記正常判定基準値と前記判定値とに基づいて正常状態から乖離した異常状態を検出することにより前記入力データの異常検出を行う異常検出手段と、を備えることを特徴としている。

【0008】この発明では、予め収集した正常データをもとに、例えば自己組織化マップを用いる等、教師なし学習を行うニューラルネットワークを用いることによって、正常データの特徴が抽出され、これに基づき正常状態モデルが形成されて記憶される。そして、正常なデータであるのか異常なデータであるのか不明な未知の入力データと正常状態モデルとに基づいて、入力データを正常とみなすことの可能な正常判定基準値が検出されると共に、入力データの異常判定のための判定値が算出される。そして正常判定基準値と判定値とに基づいて正常状態から乖離した異常状態を検出することによって、入力データの異常検出が行われる。したがって、正常状態モデルに基づいて入力データに応じた正常判定基準値が動的に設定されこれに基づき異常検出が行われるから、入力データの異常を動的に判断することが可能となる。

【0009】

【発明の実施の形態】以下に、本発明の実施の形態を説明する。図1は、本発明によるデータ異常検出装置の一例を示したものであり、このデータ異常検出装置100は、バッチプラント、連続系プラント等のデータ異常検出対象の対象プラントPが正常状態であるときに収集した、例えば対象プラントPに設けたセンサからの検出信号など、予め設定された観測用のデータを記憶するデータベース等の観測値格納装置1と、観測値格納装置1に格納された正常データをもとに自己組織化マップを用いて正常データの特徴を抽出し、これに基づいて正常状態モデルを生成しこれを正常状態モデル格納装置3に格納する特徴抽出装置2と、正常状態モデル格納装置3に格

納された正常状態モデルをもとに対象プラントPの所定の監視用データの異常検出を行う異常検知装置4と、異常検知装置4で異常を検出したときに異常表示を行うためのCRTディスプレイ等の表示装置5とから構成されている。

【0010】前記特徴抽出装置2は、マイクロコンピュータ等で構成され、例えばコホーネンの自己組織化マップ等による教師なし学習を行うニューラルネットワークを用いて、観測用データのみに基づいて内部構造を外界に適應させる自己組織化処理を行う。そしてこの処理によって形成された出力層において、多次元の監視用データがどの出力ユニットに最大反応を起こすかを調べて各出力ユニットの特徴付けを行う。

【0011】自己組織化マップは、基本的には図2に示すように、入力層と出力層との二層構造ネットワークであり、入力層と出力層とが重みベクトル w によって結合されるようになっている。特徴抽出装置2では、図3に

$$d_i = \sum [x_i(t) - w_{ij}(t)]^2 \quad \dots\dots (1)$$

$i=1$ 次に、ステップS4で、重みベクトル w_j と入力ベクトル x との距離 d_j が最小となる出力ユニットを選択し、これを勝者ユニット n_{j*} とする。

【0015】次に、ステップS5で、次式(2)にした

$$w_{ij}(t+1) = w_{ij}(t) + \eta(t) [x_i(t) - w_{ij}(t)] \quad \dots\dots (2)$$

なお、前記近傍関数 N_{j*} で定まる近傍領域は、図4に示すように、学習の初期は広く、学習の進行と共に狭くなっていくようになっている。また、式(2)中の $\eta(t)$ は、近傍関数 $N_{j*}(t)$ で指定される近傍領域内の出力ユニットに関して距離を検出する場合には $\eta(t) = \alpha(t)$ 、近傍関数 $N_{j*}(t)$ で指定される近傍領域外の出力ユニットに関して距離を検出の場合には $\eta(t) = 0$ となる値であり、つまり、近傍関数 $N_{j*}(t)$ で指定される近傍領域外の出力ユニットに関してはその重みベクトルの更新を行わないようになっている。また、前記 $\alpha(t)$ ($0 < \alpha(t) < 1$)は、学習係数であって学習の進行と共に単調減少させるようになっている。

【0017】そして、観測値格納装置1に格納されている正常データに対して所定の処理が終了したとき(ステップS6)、処理を終了する。これによって、教師なし学習により正常データの特徴を表す自己組織化マップが形成される。この自己組織化マップでは、勝者ユニット n_{j*} 及び近傍関数 N_{j*} により特定される近傍ユニットは全てその時の入力ベクトルに近づくようになっている。そして、学習の初期は、近傍関数 N_{j*} によって多くの出力ユニットが近傍ユニットと見なされて大まかなマップが形成されるが、学習の進行と共に、近傍関数 N_{j*} により勝者ユニット n_{j*} の近傍ユニットとみなされる出力ユニットが減少し、局所的な微調整が進んで空間的な解像

度が向上するようになっている。

【0012】すなわち、まず、ステップS1で、入力層と出力層との間の重みベクトル w の初期値を乱数を用いて小さな値に設定する。次に、ステップS2で、観測値格納装置1に格納されている正常な監視用データである正常データを入力ベクトル x ($x = (x_1, x_2, x_3, \dots, x_n)$)として入力層に入力する。

【0013】次に、ステップS3で、出力層の各出力ユニット n_j の重みベクトル w_j と入力ベクトル x との距離を計算する。出力層の j 番目の出力ユニット n_j の重みベクトル w_j と入力ベクトル x との距離 d_j は次式

(1)で与えられる。なお、式中の w_{ij} は入力層の i 番目の入力ユニットと出力層の j 番目の出力ユニットとの間の重みを表す。また、 t は離散時間座標である。

【0014】

がって、勝者ユニット n_{j*} と近傍関数 $N_{j*}(t)$ とによって定義される勝者ユニット n_{j*} の近傍の出力ユニットと入力ユニットとの間の重み w_{ij} を更新する。

【0016】

度が向上するようになっている。

【0018】さらに、特徴抽出装置2では、図3に示すアルゴリズムにしたがって学習した結果得られた正常データに対応する自己組織化マップに基づいて、各出力ユニットの重みベクトル w_j をもとに、各出力ユニット間の距離関係を算出し、出力ユニット間の距離関係を示す正常データの分布形を表す特徴マップを形成し、各出力ユニットの重みベクトル w_j と共に特徴マップを正常状態モデルとして正常状態モデル格納装置3に記憶する。図5は、図3に示す教師なし学習のアルゴリズムにしたがって、33個の正常データをもとに学習を行い、正常状態モデルとして自己組織化マップの出力ユニット間の距離を算出し、その距離関係を階調度の濃淡により表したものである。近接した重みベクトル間の平均距離が小さいときには薄い色合い、逆に平均距離が大きいならば濃い色合いを用いており、これによりクラスター間の峡谷を明白に示すようにしている。

【0019】一方、前記異常検知装置4では、稼働中の対象プラントPに設けた各種センサからの多次元の監視用データを入力し、この監視用データと正常状態モデル格納装置3に格納されている正常状態モデルとをもとに、図6に示すフローチャートにしたがって、監視用データの異常検出を行う。

【0020】すなわち、まず、ステップS11で、対象プラントPに設けた各種センサからの多次元の監視用デ

ータを異常検出のための入力データ（入力ベクトル） x （ $x = (x_1, x_2, x_3, \dots, x_n)$ ）として入力する。

【0021】次いで、ステップS12で、次式（3）にしたがって入力ベクトル x と、正常状態モデルの各出力

$$d_{ext} = \sum_{i=1}^n [x_i(t) - w_{ij-n}]^2 \quad \dots\dots (3)$$

次いで、ステップS13で、ステップS12で算出した各出力ユニット n_j 毎の外部距離 d_{ext} が最小となる出力ユニット n_j を検索し、これを勝者ユニット n_{j*} として仮定する。

【0023】次いで、ステップS14で、正常状態モデルにおいて、例えば図7に示す勝者ユニット n_{j*} の第1近傍領域内の全ての出力ユニットに対して、次式

$$d_{int} = \text{MAX} \left\{ \sum_{i=1}^n [w_{ij-n} - w_{ij-n*}]^2 \right\} \quad \dots\dots (4)$$

次いで、ステップS15で、外部距離 d_{ext} が内部距離 d_{int} より大きいかな否かを判定し、 $d_{ext} < d_{int}$ であるときにはステップS16に移行し、表示装置5に異常発生メッセージを表示する等の所定の異常検出時の処理を実行した後ステップS17に移行する。一方、ステップS15で、 $d_{ext} < d_{int}$ でないときには、そのままステップS17に移行し、例えばオペレータ等によって、異常検出処理の終了指示が行われたかな否かを判定し、終了指示が行われないときにはステップS11に戻って次の監視用データに対して上記と同様にして処理を行う。そして、ステップS17で異常検出処理の終了指示が行われたときには処理を終了する。

【0025】ここで、正常状態モデル格納装置3がモデル記憶手段に対応し、特徴抽出手段が特徴抽出装置2に対応し、内部距離 d_{int} が正常判定基準値に対応し、図6のステップS14の処理が正常判定基準値検出手段に対応し、外部距離 d_{ext} が判定値に対応し、図6のステップS12の処理が判定値演算手段に対応し、図6のステップS15の処理が異常検出手段に対応している。

【0026】次に、上記実施の形態の動作を説明する。今、対象プラントPについてその異常検出を行うものとする。予め、対象プラントPが正常に稼働している状態において、対象プラントPに設置した各種センサ等からの監視用データを収集する。そして、この収集した正常データに基づいて特徴抽出装置2において、図3のフローチャートにしたがって自己組織化マップを用いて特徴抽出を行い、これをもとに図4に示すような重みベクトル間の距離関係を表す特徴マップを形成して正常状態モデルとして正常状態モデル格納装置3に格納しておく。

【0027】そして、対象プラントPの異常検出を行う場合には、まず、対象プラントPに設けた各種センサか

ユニット n_j の重みベクトル w_{j-n} との間の距離である外部距離 d_{ext} を算出する。なお、式中の w_{ij-n} は、正常状態モデルにおける入力ユニット及び出力ユニット間の重みである。

【0022】

（4）にしたがって内部距離 d_{int} を算出する。すなわち、正常状態モデルにおいて勝者ユニット n_{j*} の重みベクトル w_{j*-n} 及びこれと隣接する出力ユニットの重みベクトル w_{j-n} との間の平均距離を演算し、出力ユニット間の平均距離のうちの最大値を内部距離 d_{int} として設定する。

【0024】

ら正常であるか異常であるのか不明な未知の監視用データを入力ベクトル x として入力し（ステップS11）、正常状態モデル格納装置3に格納している正常状態モデルと入力ベクトル x とをもとに前記式（3）に基づいて外部距離 d_{ext} を算出する。そして、この外部距離 d_{ext} が最小となる出力ユニット n を勝者ユニット n_{j*} とし、前記式（4）に基づいて、内部距離 d_{int} を算出する。そして、例えば、図7に示すように、近傍領域として勝者ユニット n_{j*} を囲む8つの出力ユニットを含む第1近傍領域を設定すると、これら8つの出力ユニットの重みベクトル w_j と勝者ユニット n_{j*} の重みベクトル w_{j*} との平均距離を算出し、この平均距離の最大値を内部距離 d_{int} として設定する。

【0028】ここで、入力ベクトル x が正常データであれば、この入力ベクトル x は正常状態モデルと同等の特徴を有することになるから、正常状態モデルの各出力ユニットの重みベクトル及び入力ベクトル x の距離である外部距離 d_{ext} が最小となる勝者ユニット n_{j*} と、これの近傍の出力ユニットとの距離関係は、図5の出力ユニット間の距離関係を表す特徴マップにおける距離関係を満足することになる。

【0029】したがって、入力ベクトル x と正常状態モデルの重みベクトルとのずれの最小値である外部距離 d_{ext} が、勝者ユニット n_{j*} とこの近傍の出力ユニットとの平均距離の最大値である内部距離 d_{int} 以下（内部距離 $d_{int} \geq$ 外部距離 d_{ext} ）であれば、入力ベクトル x は、正常状態モデルと同等の特徴を有するとみなすことができ、入力ベクトル x は正常と判断される（ステップS15）。

【0030】逆に、内部距離 $d_{int} <$ 外部距離 d_{ext} であるときには、外部距離 d_{ext} が、正常状態モデルに基づいて算出される、入力ベクトル x が正常であるとみな

すことの可能な上限値である内部距離 d_{int} を越えることになるから、つまり、正常状態モデルと同等の特徴を有していない可能性があるともみなされて、入力ベクトル x は異常の可能性があると判断される（ステップS15）。そして、表示装置5にデータに異常が発生したことが通知される。

【0031】したがって、上述のように、予め収集した各種の正常データをもとに、教師なし学習によって、正常データの特徴を表す正常状態モデルを形成してこれに基づき異常検出を行うようにしたから、複数種の監視用データ毎に異常判断基準のしきい値を設定する必要がなく、複数種の監視用データつまり多次元の監視用データである場合でもこれらを総合的に評価することができる。また、正常状態モデルを用いることによって入力される監視用データに応じて動的に判断基準値を変更することができるから、対象プラントの稼働状態に応じて的確な判断基準を設定することができ、多次元入力に関する総合的な異常検知をリアルタイムに実現することができる。

【0032】また、正常状態モデルを形成するためには、正常データさえあればよいから、例えば異常状態時の監視用データを収集することができないようなプラントについても確実に異常検出を行うことができる。

【0033】なお、上記実施の形態においては、外部距離 d_{ext} を、勝者ユニット n_{j*} と隣接する出力ユニットである第1近傍領域内の出力ユニットに関して算出するようにした場合について説明したが、これに限るものではなく、第1近傍領域よりも範囲の広い領域内の出力ユニットに対して算出するようにしてもよく、この近傍領域の範囲が広がるほど、異常検出の検出誤差の許容範囲が大きくなることになるから、例えば稼働状況等、対象プラントPに応じて近傍領域の範囲を設定すれば効果的である。

【0034】また、上記実施の形態においては、プラントの異常を検出するようにした場合について説明した

が、これに限らず、例えば計算機間での授受データの異常等を検出する場合等に適用することもできる。

【0035】ここで、上記実施の形態における、特徴抽出装置2での教師なし学習時の処理及び異常検知装置4での異常検出時の処理は、ROM、フレキシブルディスク、コンパクトディスク或いはハードディスク等の記憶媒体に電子的に格納されており、特徴抽出装置2及び異常検知装置4では、この記憶媒体からこれら処理プログラムを読み出してこれを実行するようになっている。

【0036】

【発明の効果】以上説明したように、本発明のデータ異常検出装置によれば、正常データに基づく教師なし学習により形成した正常状態モデルに基づいて入力データに応じた正常判定基準値を動的に設定し、これに基づいて異常検出を行うようにしたから、入力データの異常判断を動的に行うことができる。

【図面の簡単な説明】

【図1】本発明を適用したデータ異常検出装置の一例を示す構成図である。

【図2】自己組織化特徴マップの一例である。

【図3】コホーネンのネットワークアルゴリズムを表すフローチャートである。

【図4】近傍関数の一例を示す説明図である。

【図5】出力ユニット間の距離関係を表す特徴マップの一例である。

【図6】異常検知装置の処理手順の一例を示すフローチャートである。

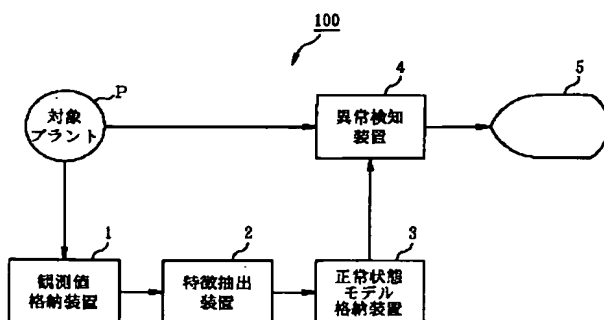
【図7】

第1近傍領域を表す説明図である。

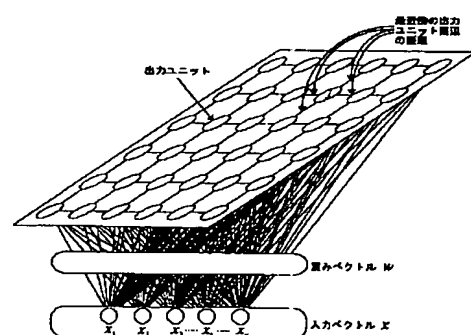
【符号の説明】

- 1 観測値格納装置
- 2 特徴抽出装置
- 3 正常状態モデル格納装置
- 4 異常検知装置
- 5 表示装置

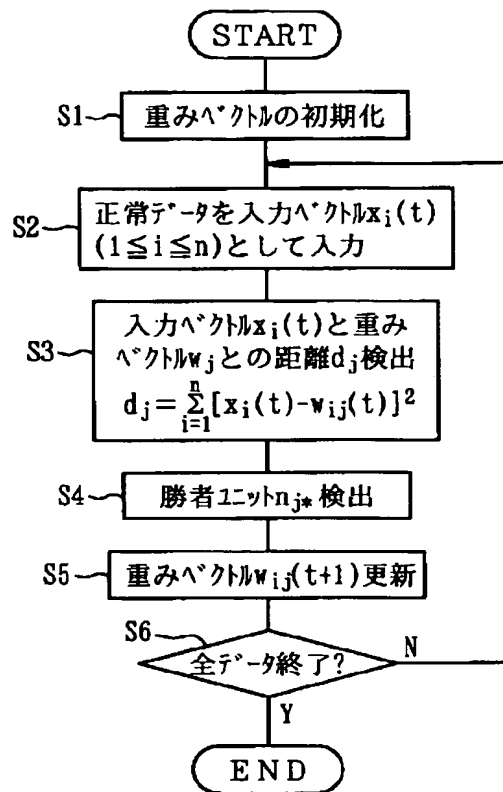
【図1】



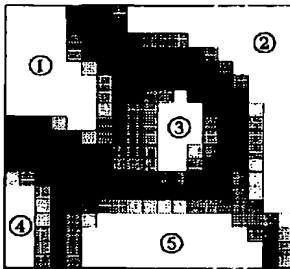
【図2】



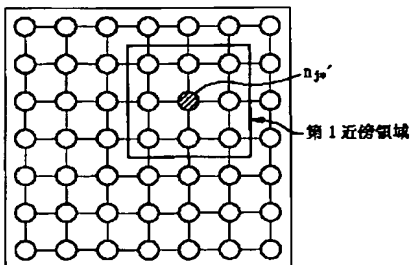
【図3】



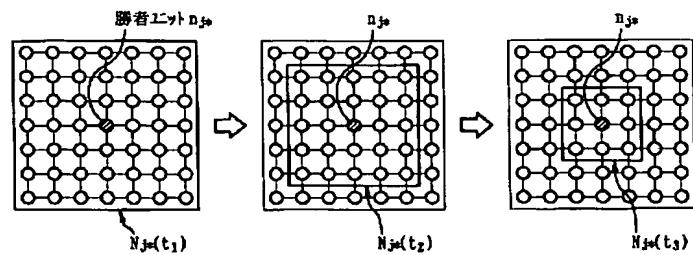
【図5】



【図7】



【図4】



【図6】

